O. ONICESCU

Member of Romanian Academy

M. C. BOTEZ

Bucharest University

A Stochastic-Informational Approach to Input-Output Analysis

Using the tools of informational statistics and informational econometrics, some random extensions of general input-output models are explored.

As particular cases, new approaches to classical models are derived.

INTRODUCTION

The mean feature in the classical econometric model focusses upon the *values* of econometric variables. But it is well-known that usually the econometric values are random variables on a (given or not) probability space: it is therefore necessary to take into account also theirs *probability distributions*. A new "space" is thus needed to be added to the classical space of econometric values, in order to describe realistically the economic behavior: the space of corresponding probability distributions — for which we propose the name of *informational space* (by analogy with O. Onicescu's *informational statistics*).

This paper deals with some properties of the informational space associated to classical input-output models.

§ 1. As an outset we will recall the general structure of input-output models. Let the elements of an n-sectorial *open* reproduction system [1] be

$$\{X_i, x_{ij}, y_i;\ i, j = 1, ..., n\}$$

where: X_i is the production of the sector i,

x_{ij} is the consumption in the sector j of the production of the sector i (i.e., the flow $i \to j$),

y_i is the (internal) consumption in the sector i of its own production;

obviously, under some conservative hypotheses, the following balance equation is fulfilled:

$$X_i = \sum_{j=1}^{n} x_{ij} + y_i, \quad i = 1, ..., n \tag{1.1}$$

Adding to the production sectors some new sectors — i.e., foreign trade, public administration, social consumption a.s.o.,— we obtain the model of a *closed* system, of the general form

$$X_i = \sum_{j=1}^{N} x_{ij} \quad i = 1, ..., N. \tag{1.2}$$

In practice, it is generally supposed that the intersectors flows are linear with respect to the productions of the receiving sectors, i.e. $x_{ij} = a_{ij} X_j$ with a_{ij} — constants. We will therefore consider the linear homogeneous system

$$X_i = \sum_{j=1}^{N} a_{ij} X_j, \quad i = 1, ..., N \tag{1.3}$$

In the classical input-output analysis it is generally supposed that $\{X_i$ x_{ij}, y_i; $i, j = 1, ..., n\}$ or $\{X_i, a_{ij}; i, j = 1, ..., N\}$ are non-random variables (i.e., deterministic analytical variables). Nevertheless for some practical purposes, it is sometimes useful to introduce the problem of (statistical) estimation of the "parameters" $\{X_i, x_{ij}, y_i; i, j = 1, ..., n\}$ or $\{X_i, a_{ij}; i, j = 1, ..., N\}$ through sampling data: in this respect, expressions like "the mean of X_i' (or x_{ij}, or y_i, or a_{ij}), or "the variance of X_i''(or x_{ij}, or y_i, or a_{ij}) become meaningfull solely if connected with the statistical inference procedures, and not with the very nature of the model. The general case in which $\{X_i, x_{ij}$, y_i; $i, j = 1, ..., n\}$ and $\{X_i, a_{ij}; i, j = 1, ..., N\}$ are random seems to have not be treated in the literature.

Our aim is to develop an approach of this type using the suggestions given by informational statistics [2] and by informational econometrics [3].

§ 2. Let then the input-output open random model be

$$\{X_i, x_{ij}, y_i; \quad i, j = 1, ..., n\}$$

where X_i, x_{ij}, y_i, $i, j = 1, ..., n$, with the economic significances given in § 1 are supposed to be non-negative random variables on a given probability space $\{\Omega, K, P\}$. Randomness can be justified by the expected fluctuations of the productions X_i, of the flows x_{ij} and of the internal consumptions y_i, due to random factors (technical progress, change of some international markets, a.s.o.). It seems therefore naturally to suppose that *a priori* we do not know the exact values of $\{X_i, x_{ij}, y_i; i, j = 1, ..., n\}$, but their probability distributions.

How such elements can be practically described? We will use some alternative approaches.

2.1. Let us suppose that $\{X_i, x_{ij}, y_i;\ i, j = 1, ..., n\}$ are non-negative (continuous) random variables, with the known distributions $\{F_i, G_{ij}, K_i;\ i, j = 1, ..., n\}$, i.e.

$$F_i(x) = \mathfrak{P}\{X_i < x\}$$

$$G_{ij}(x) = \mathfrak{P}\{x_{ij} < x\}$$

$$K_i(x) = \mathfrak{P}\{y_i < x\}$$

The balance equation (1.1) suggests that some relationships can be established between F_i, G_{ij}, K_i, $i, j = 1, ..., n$.

Let us first consider $\{x_{ij},\ j = 1, ..., n\}$ as being (statistically) independent, and also independent on y_i: that means, that the *consumptions* of the production of the sector i in the different sectors j, $j = 1, ..., n$, are independent and are independent on y_i. The random analogous of (1.1) is therefore

$$F_i(x) = \left[\bigotimes_{j=1}^{n} G_{ij}^*(x)\right] \otimes K_i(x), \quad i = 1, ..., n \tag{2.1.1}$$

by $\otimes$ meaning the classical convolution operator [1]).

A similar relation can be established among the probability densities of distributions F_i, G_{ij}, K_i (obviously, if such densities exist!). If F_i^*, G_{ij}^* and K_i^* are the characteristic functions [2]) of F_i, G_{ij} and K_i respectivelly, from (2.1.1) we obtain

$$F_i^*(t) = \left[\prod_{j=1}^{n} G_{ij}^*(t)\right] K_i^*(t), \quad i = 1, ..., n \tag{2.1.1}'$$

The relations (2.1.1) (2.1.1)' show that the random variable X_i and the sum of its parts $\left(\sum_{j=1}^{n} x_{ij} + y_i\right)$ have the same probability distribution.

But the relations (2.1.1), (2.1.1)' can be also considered as (functional) equations in (respectivelly) F_i, G_{ij} or K_i. In particular, if X_i, y_i are identically distributed, the corresponding equation seems to be similar to the well-known renewal equation; the economic interpretation of such a situation is quite trivial.

Let us now consider $x_{ij} = a_{ij}X_j$, with a_{ij} — constants. Obviously,

$$G_{ij}(x) = \mathfrak{P}\{x_{ij} < x\} = \mathfrak{P}\left\{X_j < \frac{x}{a_{ij}}\right\} = F_j\left(\frac{x}{a_{ij}}\right)$$

---

[1] $(F_1 \otimes F_2)(x) = \int_{-\infty}^{x} F_1(x - y)dF_2(y)$

[2] $F^*(t) = \int_{-\infty}^{+\infty} e^{itx}\, dF(x)$, if F is the probability distribution.

with similar modifications in (2.1.1), (2.1.1)'. But the previous hypothesis on independence of x_{ij} leads now to the independence of the sectors $\{X_i,\ i = 1, ..., n\}$. Is such an approach a realistic one? If the answer is positive, we can introduce a relation of closeness, of the form

$$\bigotimes_{i=1}^{n} F_i(x) = L(x), \tag{2.1.2}$$

where $L(x)$ is the probability distribution of the sum $\sum_{j=1}^{n} X_j$, i.e.

$$L(x) = \mathcal{P}\left\{ \sum_{j=1}^{n} X_j < x \right\},$$

with obvious forms for probability densities (if they exist!) and characteristic functions. Obviously, different other relationships between $L(x)$ and $\{G_{ij}(x),\ K_i(x);\ i, j = 1, ..., n\}$ — corresponding to alternative closeness hypotheses of the economic system — can be also considered.

Therefore, the functional relationships

$$\begin{cases} F_i(x) = \left[\bigotimes_{j=1}^{n} G_{ij}(x) \right] \otimes K_i(x), \quad i = 1, ..., n \\ \bigotimes_{i=1}^{n} F_i(x) = L(x) \end{cases}$$

define the random analogous of the classical input-output equations (1.1).

Let us finally mention that, if the random variables $\{X_i, x_{ij}, y_i;\ i, j = 1, ..., n\}$ have the means (i.e., expectations) $\{M_i, m_{ij}, n_i;\ i, j = 1, .., n\}$ respectively, from (2.1.1) it follows

$$M_i = \sum_{j=1}^{n} m_{ij} + n_i, \quad i = 1, ..., n \tag{2.1.3},$$

and further that, if $x_{ij} = a_{ij} X_i$, with a_{ij} — constants, then $m_{ij} = a_{ij} M_j$ and from (2.1.3) we obtain

$$M_i = \sum_{j=1}^{n} a_{ij} M_j + n_i, \quad i = 1, ..., n, \tag{2.1.3)'}$$

i.e. relationships of the classical input-output forms (1.1), (1.2) — with obvious analogous for closed models of the form (2.1.2).

It follows from the above considerations that our random approach is an *extension* of the classical input-output approach, and that this last one can be interpreted as leading to certain relationships between the expectations of the random elements involved in our ratio-situations.

Moreover, this random extension offers some additional advantages, Thus, if $X_i, x_{ij}, y_i,\ i, j = 1, ..., n$ have the variances $D_i, d_{ij}, e_i,\ i, j = 1, ..., n$. then, from (2.1.1) it follows (if x_{ij}, y_i are independent)

$$D_i = \sum_{j=1}^{n} d_{ij} + e_i, \quad i = 1, ..., n.$$

Therefore, using the probability distributions F_i, G_{ij}, K_i, $i, j = 1, ..., n$, we can also calculate the scattering arround the means occuring in the classical deterministic models. Obviously, the relations (2.1.3), (2.1.4) are deterministic input-output-equations too; they can be analysed using classical tools.

For practical purposes, one can estimate the probability distributions $\{F_i, G_{ij}, K_i, i, j = 1, ..., n\}$through current statistical inference tools (frequence analysis, nonparametric and parametric tests for goodness of fit a.s.o).

2.2. Let us now consider the case in which the distributions $\{G_{ij}, K_i; i, j = 1, ..., n\}$ corresponding to intersector flows $\{x_{ij}, i, j = 1, ..., n\}$ and internal sector consumptions $\{y_i, i = 1, ..., n\}$ respectivelly are not related to convolution operators, as in the case of independence studied above, suggested by (1.1). That implies a possibly more realistical assumption, that x_{ij} are dependent and that they may be correlated to y_i.

Let us replace (2.1.1) by the more general relationships

$$F_i(x) = \Phi_i[F_j(x), \quad j = 1, ..., n; \quad K_i(x)], \quad i = 1, ..., n \tag{2.2.1}$$

where Φ_i describes the way in which the probability distributions F_j, $j = 1, ..., n$ of the various sector outputs and of the internal consumption K_i are (generally) interracting so as to produce the probability distribution of the sector i.

Let us first consider Φ_i a linear mixture of the probability distributions $\{F_j, j = 1, ..., n; K_i\}$, i.e.

$$F_i(x) = \sum_{j=1}^{n} \alpha_{ij} F_j(x) + \beta_i K_i(x), \quad i = 1, ..., n, \tag{2.2.2}$$

with the additional obvious conditions

$$\alpha_{ij} \geqslant 0, \quad \beta_i \geqslant 0, \quad i, j = 1, ..., n$$

$$\sum_{j=1}^{n} \alpha_{ij} + \beta_i = 1, \quad i = 1, ..., n$$

and analogous forms for probability densities and characteristic functions.

Let us now suppose (in order to simplify the notations) that X_i, y_i are discrete random variables with the same finite set of states $\{a_1, ..., a_m\}$, and let

$$p_k^{(i)} = \mathfrak{P}\{X_i = a_k\} \quad \left(p_k^{(i)} \geqslant 0, \sum_{k=1}^{m} p_k^{(i)} = 1\right)$$

$$r_k^{(i)} = \mathfrak{P}\{y_i = a_k\} \quad \left(r_k^{(i)} \geqslant 0, \sum_{k=1}^{m} r_k^{(i)} = 1\right)$$

From (2.2.2) we obtain

$$p_k^{(i)} = \sum_{j=1}^{n} \alpha_{ij} p_k^{(j)} + \beta_i r_k^{(i)}, \quad k = 1, ..., m; \quad i = 1, ..., n. \tag{2.2.3}$$

Let us now calculate the *informational correlation* ξ_{il} of the random variables X_i, X_l; by definition,

$$\xi_{il} = \sum_{k=1}^{m} p_k^{(i)} \; p_k^{(l)};$$

the *informational auto-correlation (or self-correlation)* ξ_{ij} is known as the *informational energy* [4], and will be noted by $\varepsilon_{(i)}$

From (2.2.3) we obtain obviously

$$\xi_{li} = \sum_{j=1}^{n} \alpha_{ij} \xi_{lj} + \beta_i \eta_{li}, \quad i, l = 1, ..., n, \tag{2.2.4}$$

where η_{li} is the informational correlation of the variables X_l, y_j. Due to the obvious relation $\xi_{li} = \xi_{il}$, the following closeness equation is obtained:

$$\sum_{j=1}^{n} \alpha_{ij} \xi_{lj} + \beta_i \eta_{li} = \sum_{j=1}^{n} \alpha_{li} \xi_{ij} + \beta_l \eta_{il} \tag{2.2.5}$$

for every $i, l = 1, ..., n$.

From (2.2.4), for $l = i$, we get

$$\varepsilon_{(i)} = \sum_{j=1}^{n} \alpha_{ij} \xi_{ij} + \beta_i \eta_{ii} \quad i = 1, ..., n. \tag{2.2.6}$$

We can therefore formulate the following

Theorem: For input-output random models of the form

$$p_k^{(i)} = \sum_{j=1}^{n} \alpha_{ij} p_k^{(j)} + \beta_i r_k^{(i)} \quad k = 1, ..., m; \quad i = 1, ..., n,$$

with $p_k^{(i)} = \mathfrak{P}\{X_i = a_k\}$, $r_k^{(i)} = \mathfrak{P}\{y_i = a_k\}$, $k = 1, ..., m$; $i = 1, ..., n$, the informational energy of every X_i is a linear (convex) combination of the informational correlations of X_i with all X_j, $j = 1, ..., n$ and y_i.

Obviously, this conclusion is also true for models of a more general type (2.2.2).

A new informational interpretation of the classical input-output matrices can be now stated. Formally, one can consider every deterministic classical input-output model as being a model involving the mean values of a random input-output model, particulary of the form (2.2.2), (2.2.3). In fact, if M_i is the mean of X_i and n_i is the mean of y_i, from (2.2.2) we obtain

$$M_i = \sum_{j=1}^{n} \alpha_{ij} M_j + n_i, \quad i = 1, ..., n, \tag{2.2.7}$$

i.e. a classical input-output model. Conversely, through a Monte-Carlo type procedure one can associate to every deterministic model a random one, of the form (2.2.7). Let us remark that the deterministic models and the corresponding random models have the same input-output matrix $\| \alpha_{ij} \|$. It follows

that *the classical Leontieff input-output matrix is also the matrix of the relationships between the informational energies and correlations of the random variables, whose expectations are involved in the deterministic approach.*

Let us also mention that these results are not depending on the particular forms of the probability distributions involved in (2.2.2), (2.2.3).

The general case (2.2.1) of a nonlinear Φ_i (in the arguments F_j) is still open. If the "arguments" are separated, i.e. if the model is

$$F_i(x) = \sum_{j=1}^{n} \Phi_{ij}[F_j(x)] + \psi_i[K_i(x)], \quad i = 1,\dots, n, \tag{2.2.8}$$

with known $\Phi_{ij}[F_j(x)]$, $\psi_i[K_i(x)]$, under some restrictions, an approach of the type described above can be used.

2.3. Let us now explore the case in which the random structure of the model, i.e. of the variables $\{X_i, x_{ij}, y_i;\ i, j = 1, ..., n\}$ is no more described through the probability distributions of the individual variables, as above, but by conditional distributions of the endogeneous variables $\{X_i, i = 1, ..., n\}$, the values of the exogeneous variables $\{x_{ij}, y_i;\ i, j = 1, ..., n\}$ being given.

A random model of this kind has the form

$$\begin{aligned} &\mathcal{P}\{X_i < x |\ x_{ij},\ j = 1, ..., n;\ \ y_i\} = \\ &= F_i(x|x_{ij},\ j = 1, ..., n;\ y_i), \qquad i = 1, ..., n. \end{aligned} \tag{2.3.1}$$

A direct random extension of (1.1) suggests the model

$$F_i(x|x_{ij}\ \ j = 1,\ ...,\ n; y_i) = f_i\left(x - \sum_{j=1}^{n} x_{ij} - y_i\right), \quad i = 1, ..., n. \tag{2.3.2}$$

For $x_{ij} = a_{ij}X_j$, from (2.3.1), (2.3.2) we obtain

$$\begin{aligned} &\mathcal{P}\{X_i < x |\ X_j,\ j = 1, ..., n;\ y_i\} = \\ &= F_i(x|\ X_j,\ j = 1, ..., n;\ y_i), \quad i = 1, ..., n \end{aligned} \tag{2.3.3}$$

and respectively,

$$F_i(x|\ X_j, j = 1, ..., n;\ y_i) = f_i\left(x - \sum_{j=1}^{n} a_{ij}X_j - y_i\right) i = 1, ..., \text{n}. \tag{2.3.4}$$

Obviously, the regression analysis may be considered as a methodological tool appropiate when exploring such models. In fact, the expectations of the conditional probabilities are, by definition, the regression functions of some endogeneous variables upon the exogeneous ones.

2.4. Finally, let us consider the domain Δ in the $(n^2 + 2n)$ — dimensional euclidean space of the variables $\{X_i, x_{ij}, y_i;\ i, j = 1, ..., n\}$ verifying the relationships

$$\begin{cases} X_i \geqslant 0, \quad x_{ij} \geqslant 0, \quad y_i \geqslant 0, \quad i, j = 1, ..., n \\ X_i = \sum_{j=1}^{n} x_{ij} + y_i, \quad i = 1, ..., n. \end{cases} \tag{2.4.1}$$

A random approach to input-output analysis can be obtained introducing a probability measure $\mathfrak{P}$ on Δ: intuitively, that means to introduce a probability measure on the set of deterministic input-output models of the *same dimension n. The expression*

$$\Pi(A) = \mathfrak{P}\{(X_1, ..., X_n;\ x_{11}, ..., x_{nn}:\ \ y_1, ..., y_n) \in A\},\ A \subseteq \Delta$$

get the following meaning: $\Pi(A)$ is the probability that the variables belong to a set A in their range of variation, under the condition that every realisation satisfies the equations (2.4.1).

Obviously, for $x_{ij} = a_{ij}X_j$, $i, j = 1, ..., n$, the dimension of Δ is reduced.

For closed systems, with a_{ij} — constants, the informational energies of the outputs X_i, as well as their informational correlation, can be easily calculated. For the general case, the probability distributions of $\{x_{ij}, y_i\}$, conditionned by X_i, can be also calculated: that means, the random inputs conditionned by outputs, in a random feed-back oriented type approach [5].

§ 3. If the random variables $\{X_i, x_{ij}, y_i;\ i, j = 1, ..., n\}$ are depending on time — i.e., we are dealing with random dynamic input-output balances —, the corresponding random functional relationships can be studied using the methodological tools of multi-variate random processes. The field of research opening here seems to be quite promissing.

REFERENCES

1. LEONTIEFF, W.: *Input-Output Economics.* New York, Oxford University Press, 1966.
2. ONICESCU, O.: *Statistica Informațională* (mss) (Informational Statistique).
3. BOTEZ, M. C.: *Toward an Informational Econometrics.* Rev. Roum. Math. Pures et Appl., XXII, 10, 1381—1386, 1977.
4. ONICESCU, O.: *Energia Informațională* (Informational Energy). St. Cerc. Mat., 18, 1419—1420, 1966.
5. MANEA MĂNESCU: *Cibernetica Economică.* (Economic Cybernetics) Curs ASE, 1977 (lito).

UNE ETUDE PROBABILISTIQUE-INFORMATIONNELLE DES BALANCES ECONOMIQUES INPUT-OUTPUT

Utilisant l'appareil de la statistique et de l'économétrie informationnelles, on étudie certaines propriétés des modèles input-output pour le cas général où les variables qui s'y trouvent sont aléatoires, et on obtient, comme étant des cas particuliers, quelques interprétations nouvelles des modèles classiques.

СТОХАСТИЧНО-ИНФОРМАЦИОННАЯ ТРАКТОВКА ЭКОНОМИЧЕСКИХ БАЛАНСОВ ТИПА ВХОД — ВЫХОД

Используя аппарат информационной статистики и эконометрии, в работе рассматриваются некоторые свойства моделей типа вход-выход для общего случая, когда переменные этих моделей являются стохастичными, и получаются, в качестве частных случаев некоторые новые истолкования классических моделей.

11. IV. 1978